

■ VUFIND: EBOOKS ON DEMAND SUCHMASCHINE

von Silvia Gstrein

Zusammenfassung: eBooks on Demand (EOD) ist ein europaweites Netzwerk von mehr als 30 Bibliotheken in 12 europäischen Ländern mit dem Ziel, urheberrechtsfreie Bücher auf Anfrage zu digitalisieren und zur Verfügung zu stellen. Zur Digitalisierung kann jedes Buch bestellt werden, das im Online-Katalog der Bibliothek mit dem sog. EOD-Button versehen ist. Das so bestellte Buch wird dann innerhalb weniger Tage hochauflösend gescannt und nach erfolgreicher Bezahlung als PDF mit hinterlegtem OCR-Text zur Verfügung gestellt. Bisher musste in jedem einzelnen Katalog jeder teilnehmenden Bibliothek separat gesucht werden, um ein bestimmtes Buch zu finden. Seit Ende 2010 wird nun unter der Adresse <http://search.books2ebooks.eu> eine bibliotheksübergreifende Suchmaschine angeboten, die mit der Open Source Software VuFind realisiert wurde. Derzeit werden hier 1,8 Mio. Datensätze von 12 Bibliotheken durchsuchbar gemacht. Den NutzerInnen der bibliotheksübergreifenden Suchmaschine wird so schnell und unkompliziert Zugang zu bereits digitalisierten Werken als auch Büchern, die zur Digitalisierung auf Anfrage zur Verfügung stehen, gegeben.

Schlagwörter: EOD, eBooks on Demand, Suchmaschine, Open Source, VuFind

VUFIND: EBOOKS ON DEMAND SEARCH ENGINE

Abstract: The eBooks-on-Demand (EOD) network provides a trans-European digital document delivery service for end-users from all over the world offering digitization on demand of the public domain holdings of participating libraries. Currently the EOD network comprises over 30 libraries from 12 European countries. The service is offered through „EOD buttons“ placed at each record in the different library catalogues. Already in the first years of the service a common interface to search for all books offered for the digitization service was demanded over and over. After a process of evaluation in 2009, it was decided on implementing an instance of the open source software VuFind and ingesting records from the participating libraries. At the end of 2010, the common search interface went online under <http://search.books2ebooks.eu> and makes currently available over 1.8 Mio. records from 12 libraries – giving direct access to already digitized items as well as books offered for digitisation on demand.

Keywords: EOD, eBooks on Demand, search engine, Open Source, VuFind

Mehr als 30 Bibliotheken aus 12 europäischen Ländern nehmen derzeit an dem Netzwerk „eBooks on Demand“ teil, das von der ULB Tirol koordiniert wird. Mit diesem Service werden gemeinfreie Bücher auf Wunsch digitalisiert und an Kunden aus aller Welt verschickt.

Der eigentliche Bestellvorgang wird über einen „EOD-Button“ beim entsprechenden Datensatz im Bibliothekskatalog der jeweiligen Bibliothek ausgelöst. Das ist für die Benutzer einer Bibliothek eine sehr bequeme Möglichkeit, doch Kunden, die an einem ganz bestimmten Buch interessiert sind, mussten bisher jeden einzelnen der Kataloge durchsuchen, um herauszufinden, ob ein bestimmtes Buch auch tatsächlich über den EOD Service geliefert werden kann. Zu Recht wurde daher schon sehr früh die Forderung nach einem zentralen Einstiegspunkt für alle angebotenen Bestände gestellt.

Diese zentrale Einstiegsstelle sollte zumindest folgende Punkte abdecken:

- eine rasche Suche über alle Kataloge bzw. die für EOD relevanten Records der beteiligten Bibliotheken ermöglichen,
- den Bestellbutton problemlos beim einzelnen Datensatz einbinden,
- die Möglichkeit auf den Herkunftskatalog des Datensatzes zurückzulinken und
- eine multilinguale Oberfläche sowie die technische Möglichkeiten für die Einbindung weiterer Sprachen bieten.
- Schließlich sollte das interne Metadatenformat im europäischen Raum bekannt und gängig sein, damit möglichst wenige Transformationen beim Einlesen der Datensätze aus anderen Bibliotheken durchgeführt werden müssen.

Nach dem Einholen von Angeboten kommerzieller Anbieter war schnell klar, dass wir auf alternative Lösungen angewiesen waren, da die geforderten Summen die Kapazitäten des Netzwerks bei weitem übersteigen würden. Dies um so mehr als dieser zentrale Suchservice nur ein Service unter vielen darstellt und den eigentlichen Digitalisierung-on-Demand Service nur unterstützen sollte.

Im Laufe des Jahres 2009 haben wir mehrere Open Source Produkte für Suchmaschinen sowie OPACs evaluiert, darunter fanden sich:

- VuFind (<http://vufind.org/>),
- Blacklight (<http://projectblacklight.org/>),
- LibraryFind (<http://libraryfind.org/>),
- Scriblio (<http://about.scriblio.net/>) und
- SOPAC (<http://thesocialopac.net/>).

In die engere Auswahl kamen VuFind und Blacklight. Wir entschieden uns dann schlussendlich für VuFind aus einem sehr praktischen Beweggrund heraus: Unsere Entwickler haben schon viel Erfahrung mit den in VuFind verwendeten Technologien (PHP, Java) im Gegensatz zu den in Blacklight (Ruby on Rails). Auch für die nähere Zukunft gehen wir davon aus, dass wir leichter Programmierer mit Erfahrung in PHP und Java finden werden.

VuFind wurde ursprünglich an der Villanova University in den USA entwickelt – mit Apache Lucene und Solr als Grundlage. Alle Funktionen von Solr wie Ranking, Boosting, Stemming, Facettierung, Filterung, etc. stehen in VuFind zur Verfügung. Der Kern von VuFind ist wie bereits erwähnt in PHP programmiert und verwendet Smarty als Template Framework für die Oberfläche, die mit HTML- und CSS-Kenntnissen leicht anpassbar ist.

Ein Überblick über weltweite VuFind Installationen findet sich unter http://vufind.org/wiki/installation_status. Im deutschsprachigen Sprachraum sind vor allem BASE (Bielefeld Academic Search Engine), die Suchmaschine des GBV sowie seit neuestem das VuFind Projekt Sachsen bekannt.

Ende 2009 wurde dann mit der ersten Implementierungsphase gestartet. Zum einen überzeugte der sehr saubere Code des Open Source Projekts und zum anderen auch die schnell beantworteten Anfragen durch die weltweite Entwickler- und Anwendergemeinschaft. Zeitgleich forderten wir auch die an EOD beteiligten Bibliotheken auf, uns ihre Exportformate sowie Exportschnittstellen bekanntzugeben und Test-Datensets zur Verfügung zu stellen.

Anfang Oktober 2010 konnten wir dann in einem Projektmeeting den EOD-Partnern die erste VuFind-Installation mit den ersten eingelesenen Datensätzen präsentieren. In einer anschließenden Feedback-Runde wurden Verbesserungsvorschläge eingeholt, die bis Ende des Jahres 2010 eingebaut wurden. Damit konnte dann unsere EOD Suchmaschine erstmals live online gehen unter <http://search.books2ebooks.eu>.

Zum heutigen Zeitpunkt (Juni 2011) sind 1,8 Millionen Datensätze aus 12 Bibliotheken über die EOD Suche auffindbar. Das Abholen der Records aus den einzelnen Bibliothekskatalogen erfolgt entweder über OAI-PMH (bevorzugt) oder über Batch-Upload auf FTP. Für das Einbinden von MAB-Datensätzen wurde ein einfaches Stylesheet für die MAB-MARCXML Transformation gebaut, womit nun auch die Bücher der ULB Tirol, die für den EOD Service zur Verfügung stehen, über search.books2ebooks.eu gefunden werden können.

Die Nutzungszahlen klettern seit Inbetriebnahme stetig nach oben – derzeit nutzen durchschnittlich 600 individuelle Besucher pro Tag die Seite

– mit Spitzenwerten bis fast 1.000 Nutzer pro Tag. Über eigene XML-Sitemaps sind die Datensätze sehr gut über eine Google-Suche zu finden und dort im Ranking weit oben. Dies spiegelt sich auch im Fakt wieder, dass fast 90% der Nutzer ihren Weg über eine Google-Suche zur EOD Suchmaschine fanden.

Auch bei den steigenden Bestellzahlen im eBooks on Demand Service selbst zeigt sich die bessere Auffindbarkeit der Bücher deutlich. Ein Beispiel aus einer teilnehmenden Partnerbibliothek: Die Bayerische Staatsbibliothek hat seit dem Einspielen ihrer Datensätze in die Suchmaschine insgesamt deutlich mehr Bestellungen, die sich schließlich aus folgenden Quellen zusammensetzen: ein Drittel der Bestellungen werden über die EOD Suchmaschine abgesetzt, ein weiteres Drittel der Bestellungen erfolgt über die Test-Implementierung des EOD-Buttons in Worldcat und das restliche Drittel der Bestellungen verdankt ihren Ursprung dem lokalen OPAC der BSB bzw. dem Bayerischen Verbundkatalog sowie manueller Eingabe.

In der nächsten Zeit möchten wir natürlich primär weitere Datensätze aus noch nicht eingespielten Katalogen der weiteren EOD-Bibliotheken einspielen. Parallel dazu arbeiten wir an einem Updatemechanismus der bereits eingelesenen Datensätze sowie der Übersetzung der Oberfläche in alle im Netzwerk unterstützten Sprachen. Im Laufe des nächsten Jahres wollen wir dann die direkte Anzeige der Kosten für die Digitalisierung in der Oberfläche beim Datensatz realisieren. Schlussendlich soll search.books2ebooks.eu nicht nur Sucheinstieg für Bücher sein, die für den EOD-Service zur Verfügung stehen, sondern auch mittelfristig ein Nachweisinstrument für bereits digitalisierte Werke werden. Daher planen wir, Abfragen von APIs von großen Portalen wie z.B. Europeana und Google einzubinden, um Hinweise auf bereits vorhandene Digitalisate zu geben.

MMag.^a Silvia Gstrein,
Abteilung für Digitalisierung und Elektronische Archivierung
Universitäts- und Landesbibliothek Tirol
A-6020 Innsbruck, Innrain 52
Telefon: ++43-(0)512-507-8451
Fax: ++43-(0)512-507-949 08451
E-Mail: silvia.gstrein@uibk.ac.at
Website: <http://www.uibk.ac.at/ulb/dea/>